The STAR Data Acquisition System (DAQ) And Level III Trigger (L3)

Tonko A. Ljubicic (BNL)

(on behalf of the STAR-DAQ and STAR-L3 Groups)

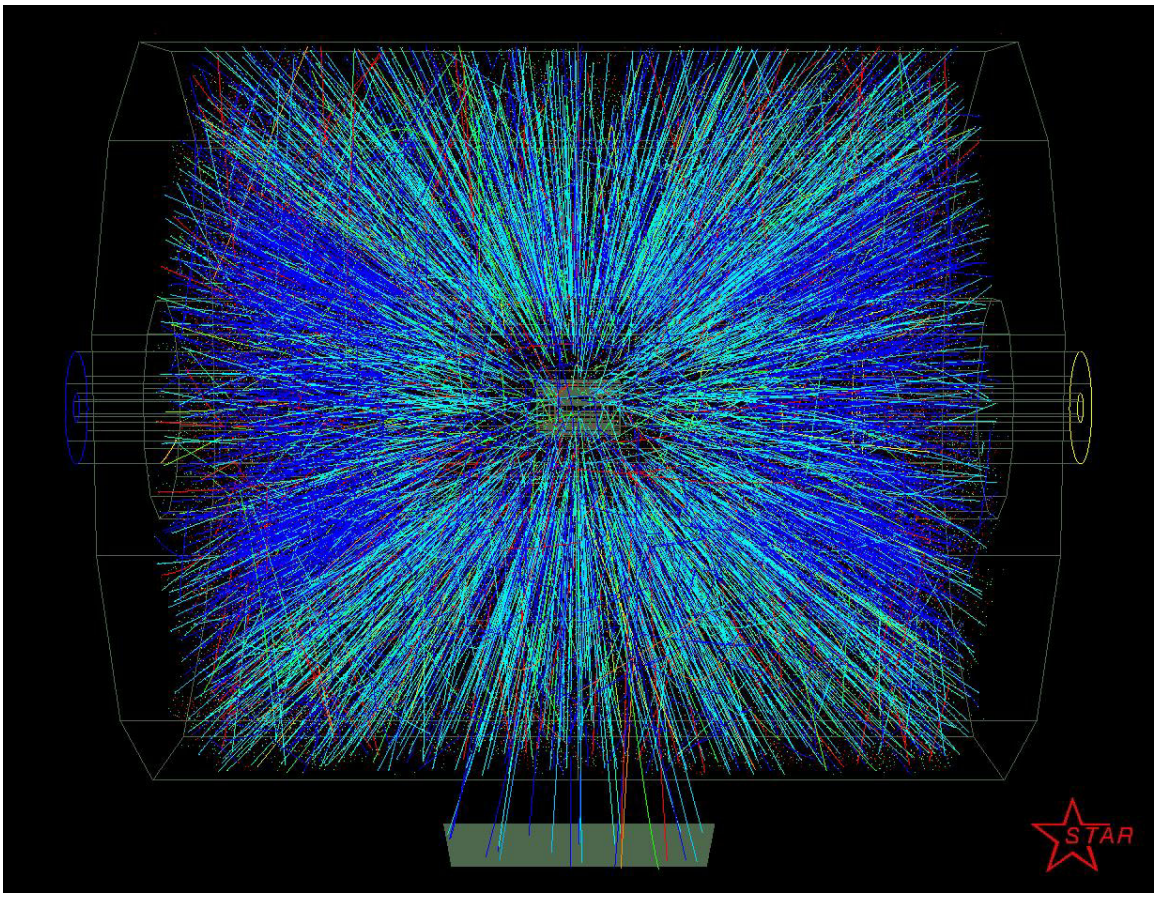
Purpose of DAQ

1. Read the *digital* data from STAR's detectors and archive them via HPSS/RCF
2. Pre-process some of the data for later use by the Level III triggering system (i.e. TPC cluster finding)

Purpose of Level III Trigger

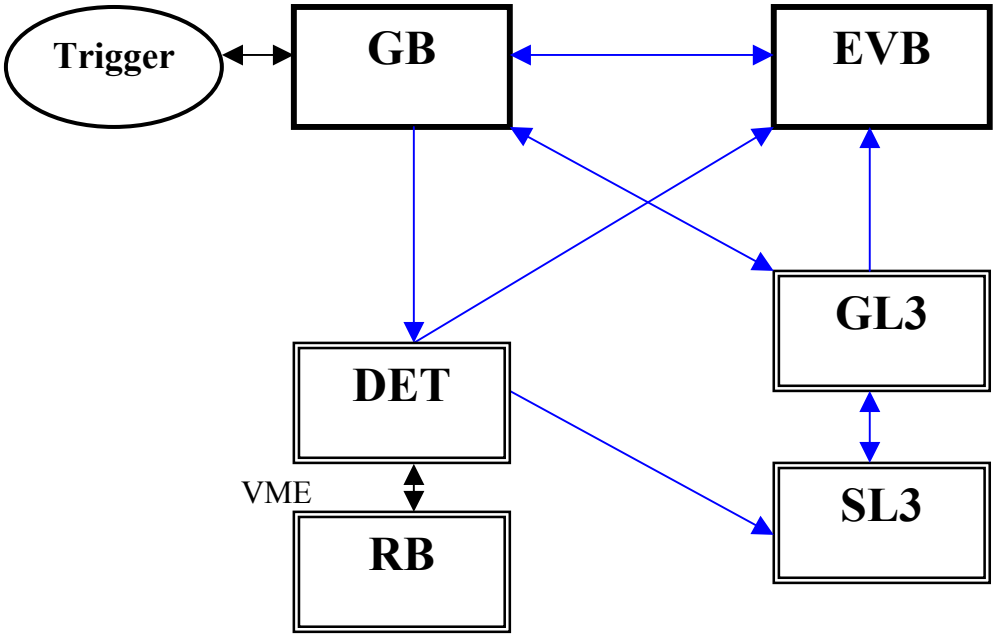
1. Perform tracking in real-time in the TPC (and other detectors) and based on the full track information (as well as any other detector data) select events deemed "more interesting" than others → rare event/particle search

Typical central Au+Au collision fully tracked and displayed in the Level III display



~8000 particle tracks; ~80 milliseconds total time

DAQ + Level III Architecture



GB Global Broker
(Motorola PPC MVME2306+vxWorks)
Organizes Trigger data and L3 decisions

EVB Event Builder
(4 CPU Sun 450+Solaris)
Gathers all detector contributions into one event and sends it to HPSS/RCF via Gigabit Ethernet TCP/IP

GL3 Global Level 3 Brokers (3)
(Linux+Alpha 21264 or Intel)
Assemble tracks from SL3s and make the final decision

SL3 Sector Level 3 CPUs (48)
(Linux+Alpha 21264@466 or 600 MHz)
Perform tracking on a TPC sector (i.e.) basis and send results to GL3

DET Detector Brokers (>20)
(Motorola PPC MVME2306+vxWorks)
RB Receiver Boards (>188)
(Custom 9U VME Boards)

- Core of DAQ.
- Housed in VME crates.
- Each RB has one optical fiber feeding it from the detector (144 TPC, 24 SVT, 20 FTPC) and 3 CPUs (Intel I960 + vxWorks) which perform cluster-finding, zero-suppression and data formatting with the help of 18 STAR-DAQ ASICs.
  Total of 564 CPUs!!!
- EMC (1) and ShowerMax (8) have no CPUs on the Receiver Board.
- RICH, TOFp & FPD have no RBs since the data is formatted by their front-end electronics (i.e. VME CRAMs or even Linux controlled CAMAC) and shipped via TCP/IP.

Connected together via Myrinet

- Fast (+128 MB/s), switched, low latency, proprietary network from Myricom Inc.
- PMC & PCI modules
- Has excellently supported software drivers for Solaris, Linux & vxWorks on PowerPC, Pentium, SPARC and Alpha architectures

Level III Alpha CPU Farm

DAQ VME Crates with Receiver Boards

For more information:

Design and Implementation of the STAR experiment's DAQ
A. Ljubicic, Jr, et al (STAR DAQ collaboration).
IEEE Trans. Nucl. Sci. 45, No. 4, p283- (1998)

The proposed level-3 trigger system for STAR.
Adler, C et al.
IEEE Trans. Nucl. Sci., Vol 47, No. 2, pp. 358-361, April 2000.

The STAR DAQ receiver board.
LeVine, M.J., Ljubicic, A., Jr., Schulz, M., Scheetz, R., Consiglio, C., Padrazo, D., and Zhao, Y.
IEEE Trans. Nucl. Sci. Vol 47, No. 2, pp. 127-131, April 2000.

The STAR Experiment's Data Acquisition System.
Ljubicic, A. Jr. et al (STAR-DAQ Collaboration).
IEEE Trans. Nucl. Sci., Vol 47, No. 2, pp. 99-102, April 2000

THE STAR LEVEL-3 TRIGGER SYSTEM.
J.S. Lange, C. Adler, J. Berger, M. Demello, D. Flierl, J. Landgraf, M.J. LeVine, A. Ljubicic, J. Nelson, D. Rohrich, J.J. Schambach, D. Schmischke, M.W. Schulz, R. Stock, C. Struck, P. Yepes (Frankfurt U. & Rice U. & Brookhaven & Birmingham U. & Bergen U. & Texas U. & Heidelberg U.).
Nucl.Instrum.Meth.A453:397-404,2000

THE IMPLEMENTATION OF THE STAR DATA ACQUISITION SYSTEM USING A MYRINET NETWORK
J.M.Landgraf, C.Adler, M.J.LeVine, A.Ljubicic Jr,J.M. Nelson, M.W.Schulz and J.S.Lange
IEEE Trans. Nuc. Sci. 48, No. 3 (2001)

(STAR DAQ & Level III papers to be published in the future NIM volume dedicated to RHIC)

At a glance…

Level3

- Cluster of parallel processing units run in a round-robin algorithm: scalable, redundant/robust
- Each unit runs specific code which needs fast FPU, not too much memory (128 MB is enough) → benefits from fast CPUs, large L2 caches, fast memory architectures
- Each node can be replaced by a faster CPU if money is available.
- Cluster can be resized at will (more money → more CPUs → more power)

DAQ

- Massively parallel system (almost 600 CPUs!) → fast cluster finding & zero-suppression.
- Hierarchically organized (sub detectors are leaves of the system, detectors are branches) → easily scales to many different detectors as long as the network withstands the rates (not a problem yet…)
- Programmable→ due to CPU (and RTOS) presence algorithms can be easily improved, modified or changed with time
- But Receiver Boards are not scalable (all boards are custom, CPUs are slow), not extendible (no more ASICs), not adaptable (boards are designed for TPC, SVT & FTPC only) → will need to be replaced by something else

Current Performance Snapshot

Aggregate input rate to DAQ:

- TPC+SVT+FTPC ~ 250 Gbits/second (100 events/sec ➔ given by front-end electronics)
- Small detectors (i.e. EMC) ~ 1000 events/sec

Level III Rates

~ 50 central events/second (i.e. 8000 tracks each)

Output of DAQ (Event Builder)

~ 60 MB/sec (about 7-8 central events/sec)

Future Possibilities

Level 3

- More CPU nodes…
- Faster CPUs…
- Even better algorithms…
- Extension to different detectors (FTPC & SVT already have the cluster finder working; other detectors (i.e. EMC) are easily incorporated)

DAQ

- Event Builder Improvements → multiple Myrinet adapters; more buffered disk storage (currently 700 GB)

- Even further → dual (multiple?) Event Builder Nodes
  - current system can scale easily by adding more parallelism (rates of 1 GB/s are possible by adding more Event Builders and going round-robin on a per event basis)

But…

- What do we do with the data of the order of 1 GB/sec ??? when already after 50 MB/s (sustained) we need HPSS upgrades
- …and who and how will read and analyze this huge amount?
- …and where do we physically store all those tapes?
- …and how do we handle them?

Some (short term) possibilities:

- Save only the 2D peak coordinates for the TPC, SVT & FTPC and not the raw data → ***saves about 5X***

- Aggressive use of Level 3 (and other triggers) to cut down the number of "vanilla" events